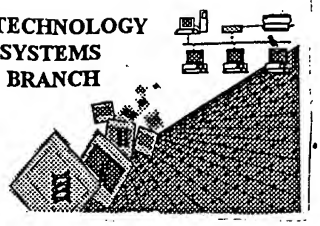
BIOTECHNOLOGY SYSTEMS BRANCH



0520/0280

RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/099,791
Source: OIPE
Date Processed by STIC: 4/2/02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>, **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. **Hand Carry directly to:**
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
   Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to:** U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

OIPE

Raw Sequence Listing Error Summary

ERROR DETECTED | SUGGESTED CORRECTION | SERIAL NUMBER: 10/099,791

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1_____Wrapped Nucleics
Wrapped Aminos
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2_____Invalid Line Length
The rules require that a line not exceed 72 characters in length. This includes white spaces.

3_____Misaligned Amino Numbering
The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4_____Non-ASCII
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5_____Variable Length
Sequence(s)____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6_____PatentIn 2.0 "bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7_____Skipped Sequences (OLD RULES)
Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8_____Skipped Sequences (NEW RULES)
Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9_____Use of n's or Xaa's (NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10_____Invalid <213> Response
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11_____Use of <220>
Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12_____PatentIn 2.0 "bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13_____Misuse of n
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

OIPE

Does Not Comply
Corrected Diskette Needed

Errors on pp. 3-5

RAW SEQUENCE LISTING DATE: 04/02/2002
PATENT APPLICATION: US/10/099,791 TIME: 15:37:47

Input Set : A:\CEN 0245 Seq.txt
Output Set: N:\CRF3\04022002\J099791.raw

The type of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.

```
      3 <110> APPLICANT: Heiskala, Marja
      5 <120> TITLE OF INVENTION: REG-LIKE PROTEIN
      7 <130> FILE REFERENCE: CEN0285
C-->  9 <140> CURRENT APPLICATION NUMBER: US/10/099,791
C-->  9 <141> CURRENT FILING DATE: 2002-03-14
      9 <160> NUMBER OF SEQ ID NOS: 45
     11 <170> SOFTWARE: PatentIn version 3.0
     13 <210> SEQ ID NO: 1
     14 <211> LENGTH: 477
     15 <212> TYPE: DNA
     16 <213> ORGANISM: human
     18 <400> SEQUENCE: 1
     19 atggcttcca gaagcatgcg gctgctccta ttgctgagct gcctggccaa aacaggagtc     60
     21 ctgggtgata tcatcatgag acccagctgt gctcctggat ggttttacca caagtccaat    120
     23 tgctatggtt acttcaggaa gctgaggaac tggtctgatg ccgagctcga gtgtcagtct    180
     25 tacggaaacg gagcccacct ggcatctatc ctgagtttaa aggaagccag caccatagca    240
     27 gagtacataa gtggctatca gagaagccag ccgatatgga ttggcctgca cgacccacag    300
     29 aagaggcagc agtggcagtg gattgatggg gccatgtatc tgtacagatc ctggtctggc    360
     31 aagtccatgg gtgggaacaa gcactgtgct gagatgagct ccaataacaa ctttttaact    420
     33 tggagcagca acgaatgcaa caagcgccaa cacttcctgt gcaagtaccg accatag       477
     36 <210> SEQ ID NO: 2
     37 <211> LENGTH: 158
     38 <212> TYPE: PRT
     39 <213> ORGANISM: human
     41 <400> SEQUENCE: 2
     43 Met Ala Ser Arg Ser Met Arg Leu Leu Leu Leu Leu Ser Cys Leu Ala
     44 1               5                  10                  15
     46 Lys Thr Gly Val Leu Gly Asp Ile Ile Met Arg Pro Ser Cys Ala Pro
     47             20                  25                  30
     49 Gly Trp Phe Tyr His Lys Ser Asn Cys Tyr Gly Tyr Phe Arg Lys Leu
     50         35                  40                  45
     52 Arg Asn Trp Ser Asp Ala Glu Leu Glu Cys Gln Ser Tyr Gly Asn Gly
     53     50                  55                  60
     55 Ala His Leu Ala Ser Ile Leu Ser Leu Lys Glu Ala Ser Thr Ile Ala
     56 65                  70                  75                  80
     58 Glu Tyr Ile Ser Gly Tyr Gln Arg Ser Gln Pro Ile Trp Ile Gly Leu
     59                 85                  90                  95
     61 His Asp Pro Gln Lys Arg Gln Gln Trp Gln Trp Ile Asp Gly Ala Met
     62             100                 105                 110
     64 Tyr Leu Tyr Arg Ser Trp Ser Gly Lys Ser Met Gly Gly Asn Lys His
     65         115                 120                 125
     67 Cys Ala Glu Met Ser Ser Asn Asn Asn Phe Leu Thr Trp Ser Ser Asn
     68     130                 135                 140
```

RAW SEQUENCE LISTING DATE: 04/02/2002
PATENT APPLICATION: US/10/099,791 TIME: 15:37:47

Input Set : A:\CEN 0245 Seq.txt
Output Set: N:\CRF3\04022002\J099791.raw

```
70 Glu Cys Asn Lys Arg Gln His Phe Leu Cys Lys Tyr Arg Pro
71 145                 150                 155
73 <210> SEQ ID NO: 3
74 <211> LENGTH: 78
75 <212> TYPE: DNA
76 <213> ORGANISM: human
78 <400> SEQUENCE: 3
79 atggcttcca gaagcatgcg gctgctccta ttgctgagct gcctggccaa aacaggagtc     60
81 ctgggtgata tcatcatg                                                  78
84 <210> SEQ ID NO: 4
85 <211> LENGTH: 26
86 <212> TYPE: PRT
87 <213> ORGANISM: human
89 <400> SEQUENCE: 4
91 Met Ala Ser Arg Ser Met Arg Leu Leu Leu Leu Leu Ser Cys Leu Ala
92 1               5                   10                  15
94 Lys Thr Gly Val Leu Gly Asp Ile Ile Met
95             20                  25
97 <210> SEQ ID NO: 5
98 <211> LENGTH: 17
99 <212> TYPE: PRT
100 <213> ORGANISM: human
102 <400> SEQUENCE: 5
104 Cys Ala Glu Met Ser Ser Asn Asn Asn Phe Leu Thr Trp Ser Ser Asn
105 1               5                   10                  15
107 Glu
110 <210> SEQ ID NO: 6
111 <211> LENGTH: 25
112 <212> TYPE: PRT
113 <213> ORGANISM: human
115 <400> SEQUENCE: 6
117 Cys Tyr Gly Tyr Phe Arg Lys Leu Arg Asn Trp Ser Asp Ala Glu Leu
118 1               5                   10                  15
119 Glu Cys Gln Ser Tyr Gly Asn Gly Ala
120             20                  25
122 <210> SEQ ID NO: 7
123 <211> LENGTH: 23
124 <212> TYPE: PRT
125 <213> ORGANISM: human
127 <400> SEQUENCE: 7
129 Trp Ile Asp Gly Ala Met Tyr Leu Tyr Arg Ser Trp Ser Gly Lys Ser
130 1               5                   10                  15
132 Met Gly Gly Asn Lys His Cys
133             20
135 <210> SEQ ID NO: 8
136 <211> LENGTH: 17
137 <212> TYPE: PRT
138 <213> ORGANISM: human
140 <400> SEQUENCE: 8
```

RAW SEQUENCE LISTING DATE: 04/02/2002
PATENT APPLICATION: US/10/099,791 TIME: 15:37:47

Input Set : A:\CEN 0245 Seq.txt
Output Set: N:\CRF3\04022002\J099791.raw

```
142 Cys Ala Glu Met Ser Ser Asn Asn Asn Phe Leu Thr Trp Ser Ser Asn
143 1               5                   10                  15
145 Glu
148 <210> SEQ ID NO: 9
149 <211> LENGTH: 29
150 <212> TYPE: PRT
151 <213> ORGANISM: human
153 <400> SEQUENCE: 9
155 Cys Ala Glu Met Ser Ser Asn Asn Asn Phe Leu Thr Trp Ser Ser Asn
156 1               5                   10                  15
158 Glu Cys Asn Lys Arg Gln His Phe Leu Cys Lys Tyr Arg
159             20                  25
161 <210> SEQ ID NO: 10
162 <211> LENGTH: 27
163 <212> TYPE: PRT
164 <213> ORGANISM: human
166 <400> SEQUENCE: 10
168 Cys Glu Tyr Ile Ser Gly Tyr Gln Arg Ser Gln Pro Ile Trp Ile Gly
169 1               5                   10                  15
171 Leu His Asp Pro Gln Lys Arg Gln Gln Trp Gln
172             20                  25
174 <210> SEQ ID NO: 11
175 <211> LENGTH: 23
176 <212> TYPE: PRT
177 <213> ORGANISM: human
179 <400> SEQUENCE: 11
181 Cys Gln Ser Tyr Gly Asn Gly Ala His Leu Ala Ser Ile Leu Ser Leu
182 1               5                   10                  15
184 Lys Glu Ala Ser Thr Ile Ala
185             20
187 <210> SEQ ID NO: 12
188 <211> LENGTH: 20
189 <212> TYPE: DNA
190 <213> ORGANISM: synthetic construct
192 <400> SEQUENCE: 12
193 cagctgtgct cctggatggt                                          20
196 <210> SEQ ID NO: 13
197 <211> LENGTH: 20
198 <212> TYPE: DNA
199 <213> ORGANISM: synthetic construct
201 <400> SEQUENCE: 13
202 tggtcggtac ttgcacagga                                          20
205 <210> SEQ ID NO: 14
206 <211> LENGTH: 20
207 <212> TYPE: DNA
208 <213> ORGANISM: synthetic construct
210 <400> SEQUENCE: 14
211 ctcctattgc tgagctgcct                                          20
214 <210> SEQ ID NO: 15
```

- invalid response, see error summary sheet item 10

RAW SEQUENCE LISTING DATE: 04/02/2002
PATENT APPLICATION: US/10/099,791 TIME: 15:37:47

Input Set : A:\CEN 0245 Seq.txt
Output Set: N:\CRF3\04022002\J099791.raw

215 <211> LENGTH: 20
216 <212> TYPE: DNA
217 <213> ORGANISM: synthetic construct
219 <400> SEQUENCE: 15
220 attcgttgct gctccaagtt 20
223 <210> SEQ ID NO: 16
224 <211> LENGTH: 19
225 <212> TYPE: DNA
226 <213> ORGANISM: synthetic construct
228 <400> SEQUENCE: 16
229 ttccagaagc atgcggctg 19
232 <210> SEQ ID NO: 17
233 <211> LENGTH: 19
234 <212> TYPE: DNA
235 <213> ORGANISM: synthetic construct
237 <400> SEQUENCE: 17
238 acaggaagtg ttggcgctt 19
241 <210> SEQ ID NO: 18
242 <211> LENGTH: 19
243 <212> TYPE: DNA
244 <213> ORGANISM: synthetic construct
246 <400> SEQUENCE: 18
247 atggcttcca gaagcatgc 19
250 <210> SEQ ID NO: 19
251 <211> LENGTH: 20
252 <212> TYPE: DNA
253 <213> ORGANISM: synthetic construct
255 <400> SEQUENCE: 19
256 ctatggtcgg tacttgcaca 20
258 <210> SEQ ID NO: 20
259 <211> LENGTH: 20
260 <212> TYPE: DNA
261 <213> ORGANISM: synthetic construct
263 <400> SEQUENCE: 20
264 cttgctctat ggtcggtact 20
267 <210> SEQ ID NO: 21
268 <211> LENGTH: 21
269 <212> TYPE: DNA
270 <213> ORGANISM: synthetic construct
272 <400> SEQUENCE: 21
273 actgggacca ctggagacac t 21
276 <210> SEQ ID NO: 22
277 <211> LENGTH: 19
278 <212> TYPE: DNA
279 <213> ORGANISM: synthetic construct
281 <400> SEQUENCE: 22
282 gagacactga agaaggcag 19
285 <210> SEQ ID NO: 23
286 <211> LENGTH: 20

RAW SEQUENCE LISTING DATE: 04/02/2002
PATENT APPLICATION: US/10/099,791 TIME: 15:37:47

Input Set : A:\CEN 0245 Seq.txt
Output Set: N:\CRF3\04022002\J099791.raw

287 <212> TYPE: DNA
288 <213> ORGANISM: synthetic construct
290 <400> SEQUENCE: 23
291 agacccagct gtttcatagg 20
294 <210> SEQ ID NO: 24
295 <211> LENGTH: 20
296 <212> TYPE: DNA
297 <213> ORGANISM: synthetic construct
299 <400> SEQUENCE: 24
300 aatggagaga gggcagaagg 20
303 <210> SEQ ID NO: 25
304 <211> LENGTH: 23
305 <212> TYPE: DNA
306 <213> ORGANISM: synthetic construct
308 <400> SEQUENCE: 25
309 tgatatcatc atgagaccca gct 23
312 <210> SEQ ID NO: 26
313 <211> LENGTH: 21
314 <212> TYPE: DNA
315 <213> ORGANISM: synthetic construct
317 <400> SEQUENCE: 26
318 agacagtcat ccatttgccc a 21
321 <210> SEQ ID NO: 27
322 <211> LENGTH: 21
323 <212> TYPE: DNA
324 <213> ORGANISM: synthetic construct
326 <400> SEQUENCE: 27
327 tgggcaaatg gatgactgtc t 21
330 <210> SEQ ID NO: 28
331 <211> LENGTH: 21
332 <212> TYPE: DNA
333 <213> ORGANISM: synthetic construct
335 <400> SEQUENCE: 28
336 ctctagaatc caacaaaact c 21
339 <210> SEQ ID NO: 29
340 <211> LENGTH: 21
341 <212> TYPE: DNA
342 <213> ORGANISM: synthetic construct
344 <400> SEQUENCE: 29
345 tgccagacca ggatctgtac a 21
348 <210> SEQ ID NO: 30
349 <211> LENGTH: 19
350 <212> TYPE: DNA
351 <213> ORGANISM: synthetic construct
353 <400> SEQUENCE: 30
354 atccatatcg gctggcttc 19
357 <210> SEQ ID NO: 31
358 <211> LENGTH: 20
359 <212> TYPE: DNA

VERIFICATION SUMMARY DATE: 04/02/2002
PATENT APPLICATION: **US/10/099,791** TIME: 15:37:48

Input Set : **A:\CEN 0245 Seq.txt**
Output Set: **N:\CRF3\04022002\J099791.raw**

L:9 M:270 C: Current Application Number differs, Replaced Current Application No
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date